MEMORANDUM

DEPARTMENT OF ENVIRONMENTAL QUALITY
WATER OPERATIONS

SUBJECT: Guidance Memo No. 96-007
Evaluation of Calibration Curve Linearity

TO: Regional Directors

FROM: Larry Lawson, P.E.

DATE: September 18, 1996

COPIES: Regional Office Permit Managers, Regional Office Compliance and Enforcement Managers, Martin Ferguson, Jack Vanderland, Amy Clark

During a recent series of laboratory audits the need for a tool to evaluate the linearity of calibration "curves" became apparent. Calibration curves are derived by plotting known standard concentrations against an instrument measured response. For colorimetric analyses the response is measured with either a spectrophotometer or colorimeter measuring the absorbance or transmittance of a specific wavelength of light. For electrometric analyses the response is measured in mV or in some cases as a direct readout as concentration. Using a calibration curve, an unknown analyte concentration can be derived from a sample by measuring it's response and then finding the corresponding y-axis intercept. Accurate results can be obtained when the relationship between concentration and response is linear. As the data becomes less linear accuracy suffers. A curve can be constructed graphically on paper or mathematically to obtain results. Intuitively, it is easy to see how "good" a curve is, when the best fit line is drawn between data points. Obtaining results graphically can be misleading however if care is not taken drawing a curve or if the data is not absolutely linear.

Mathematically a best fit line can be drawn using the linear regression or least squares analysis. Linear regression draws a best fit line that can be represented as:

$y = mx + b$

Where: (from the spectrophotometric example)

x = concentration

y = absorbance (or transmittance)

m = the slope of the line

b = y intercept

Once the slope and y intercept are known it is a simple matter to calculate an unknown concentration from sample response. Manually calculating a linear regression is cumbersome. Fortunately with scientific and statistical calculators it is a simple matter of inputing x and y values to calculate a line of best fit. Performing a linear regression doesn't indicate how tight the data is around the line of best fit. A method of statistically measuring how closely a straight line represents the data points of an xy graph is the correlation coefficient (*r*). As with the linear regression calculations it is a time consuming exercise to manually compute *r* values, but is a simple matter with a scientific calculator.

An *r* value of ± 1 indicates a perfect linear relationship between the data points and the line. Positive correlation coefficients indicate a positive slope and negative values indicate a negative slope. Table 1 lists the minimum acceptable *r* values for the corresponding number of standards used to develop a curve. *r* is to be used only with data that is linear or data that can be transformed to behave linearly; e.g. log, reciprocal, square root, cube root, or square transformations. Correlation coefficients can be used solely to accept or reject data. Instrumentation used in spectrophotometric\colorimetric and electrometric analysis should be capable of exceeding these requirements. Therefore this criteria should trigger a closer look at procedures by inspectors when it is not met.

TABLE 1

NUMBER OF STANDARDS	MINIMUM *r* FOR 99.9% PROBABILITY OF A LINEAR RELATIONSHIP
4	.999
5	.991
6	.974
7	.951
8	.925

If inspectors are not familiar with these computations the following tables contain "real" data for them to practice with. The data in Table 2 is for spectrophotometric total phosphorus analysis.

TABLE 2

Total Phosphorus		
Sample ID	Absorbance	Concentration
Blank	0.000	0.000
Standard 1	0.007	0.01
Standard 2	0.083	0.10
Standard 3	0.213	0.25
Standard 4	0.424	0.50
Standard 5	0.916	1.25
Blank Spike 0.25	0.207	102% Recovery
Sample 1	0.308	.795
Sample 1 0.25 Spike	0.501	105% Recovery
Sample 2	0.044	0.08
Sample 3	0.166	0.41
Sample 3 Duplicate	0.162	0.40
Sample 4	0.011	<0.01

The ***r*** factor for this calibration curve is 0.998. Before practicing on this data set keep in mind that there is a 2X dilution factor for the samples because 50 mL sample volumes were diluted to 100 mL. Standards were digested but not diluted. Since there were 5 standards the ***r*** value of 0.998 exceeds the minimum of 0.991. This indicates that the calibration curve can be used to determine unknown concentrations of phosphorus in effluent samples. Evaluating this data set one should be confident of the reported values because of the excellent correlation factor, and supporting QC. Spike recovery indicates good analytical technique and the lack of matrix interference. Given this type of supporting information an inspector would have latitude to overlook minor procedural discrepancies observed during an inspection.

Table 3 contains data from Ion Specific Electrode (ISE) ammonia measurements. *It is first necessary to perform a logarithmic transformation on concentration values before calculating the linear regression.* Do not use the blank value in the regression calculation because zero cannot be converted into a logarithm.

TABLE 3

ISE Ammonia		
Sample ID	mV	Concentration
Blank	+169.2	
Standard 1	+135.8	0.1
Standard 2	+97.6	0.5
Standard 3	+82.6	1.0
Standard 4	+45.2	5.0
Standard 5	+28.4	10.0
Standard 6	-24.8	100.0
Standard 7	-81.7	1000.0
Sample 1	+17.6	15.6
Sample 2	+107.3	.34
Sample 3	+97.5	.52
Sample 4	+80.8	1.1
Sample 4 Spike 5.0	+39.1	6.25 104% Recovery
Sample 5	+62.8	2.3
Sample 6	+83.7	.94
Sample 6 Duplicate	+83.5	.95
Sample 7	-24.2	92.2
Sample 8	+28.8	9.7

An r value of .9999 indicates a good xy correlation (7 standards - ***r*** value minimum is 0.951) and spikes and duplicate analysis show good recovery and precision respectively.

To illustrate the need to calculate a correlation coefficient with graphically determined data the following two data sets are offered. Table 4 contains data from a contract laboratory performing electrometric ammonia analysis.

TABLE 4

ISE Ammonia			
Sample ID	mV (x)	Concentration mg/L (from graph)	Concentration (y) mg/L (calculated)
Blank	108.7		
Standard 1	50.3	0.1	
Standard 2	90.0	0.5	
Standard 3	107.3	1.0	
Standard 4	164.2	10.0	
Standard 5	204.9	50.0	
Standard 6	222.4	100.0	
ERA QC Sample 5.0 mg/L	151.3	5.4	5.8
Standard 5.0 mg/L	146.9	4.7	4.9
Sample 1	169.3	12	11.9
Sample 2	131.5	2.5	2.6
Sample 3	169.3	11	12.0
Sample 4	170.1	12	12.4
Sample 4 dup.	169.6	12	12.2
Sample 4 Spike 5 mg/L	180.2	17	18.6
Standard 5.0 mg/L	146.9	4.7	4.9

The curve used to draw this data set can be seen in figure 1. The correlation coefficient is .99998 which exceeds 0.974 indicating a good correlation. Note the difference between obtaining analyte concentration from the curve and from a linear regression. Values are close but some error is introduced by interpolating off the graph. It is therefore suggested for inspectors to recommend analysts use linear regression rather than an actual curve to produce sample results. Most larger laboratories are currently using linear regression to calculate results because it is quicker and easier. It is not necessary to recalculate the regression for each sample run because the calculations are stored in memory. Standards and quality control samples (ERA) from a secondary source were analyzed throughout the sample run and the sample spike indicates no matrix interference.

The data in Table 5 was prepared by the same analyst as the last example. Looking at the data it appears that meter and analyst were producing good results; unfortunately this was not the case.

TABLE 5

ISE Ammonia			
Sample ID	mV (x)	Concentration mg/L (from graph)	Concentration (y) mg/L (calculated)
Blank	101.3		
Standard 1	54.7	0.1	
Standard 2	94.4	0.5	
Standard 3	111.7	1.0	
Standard 4	169.7	10.0	
Standard 5	209.2	50.0	
Standard 6	227.3	100.0	
ERA QC Sample 5.0 mg/L	156.1	5.2	7.2
Standard 5.0 mg/L	154.6	4.8	6.9
Sample 1	158.7	5.8	7.9
Sample 1 Dup.	164.5	6.4	9.6
Sample 1 Spike 5.0 mg/L	174.5	11	13.5
Standard 5.0 mg/L	156.2	5.1	7.2

The correlation coefficient for this six standard curve is 0.9693 which doesn't meet the criteria of 0.974. Based on this information the standards and samples must be rerun until the problem is determined and corrected. Rejection of the data is born out when graphed results, as shown in figure 2, are compared to results obtained from the regression equation. The ERA 5.0 mg/l standard shows an excellent recovery of 104% for the graphed result where the calculated recovery is 144%. Examination of the data shows this discrepancy consistently.

Therefore in the future inspectors should review calibration curves to insure that the minimum correlation coefficient is met. This means ***r*** must be calculated for each new

calibration curve created. Based on this guidance it will now be necessary for analysts to rerun standards until the minimum correlation coefficient for the corresponding number of standards is met Analysts should be encouraged to either use direct concentration readout instrumentation or to calculate results using the regression equation. Inspectors should keep in mind that graphed results are still acceptable and can be expected to produce good results.

Questions or comments regarding this topic can be directed to Bill Purcell at (804) 698-4048.

DISCLAIMER

This document provides technical and procedural guidance to the inspection staff to evaluate laboratories producing data related to permit compliance. This document is guidance only. It does not establish a binding norm and is not finally determinative of the issues addressed. Agency decision in any particular case will be made by applying the State Water Control Law and the implementation regulations on the basis of site specific facts.

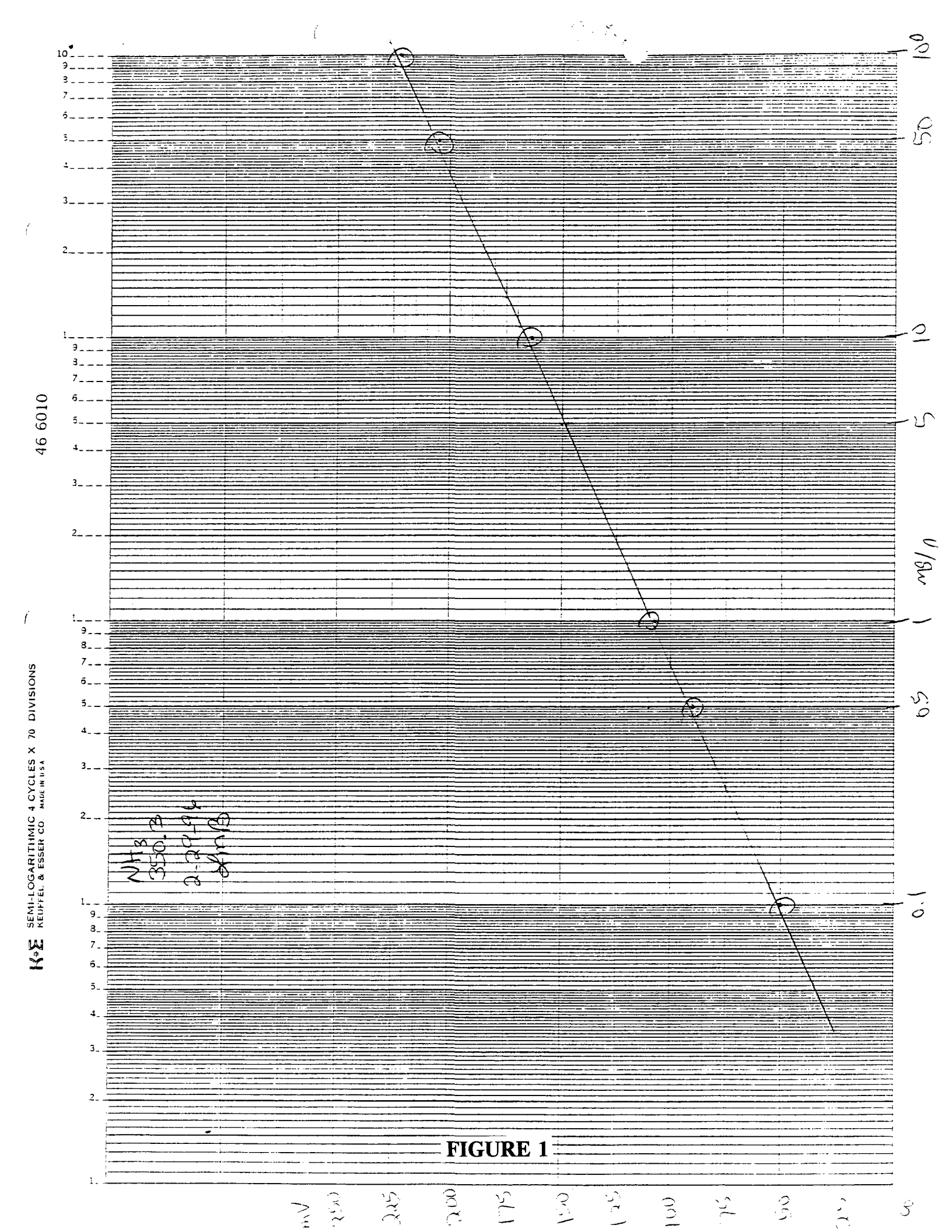
46 6010
SEMI-LOGARITHMIC 4 CYCLES X 70 DIVISIONS
KEUFFEL & ESSER CO. MADE IN U.S.A.
mg/l
100
50
10
5
1
0.5
0.1

FIGURE 1

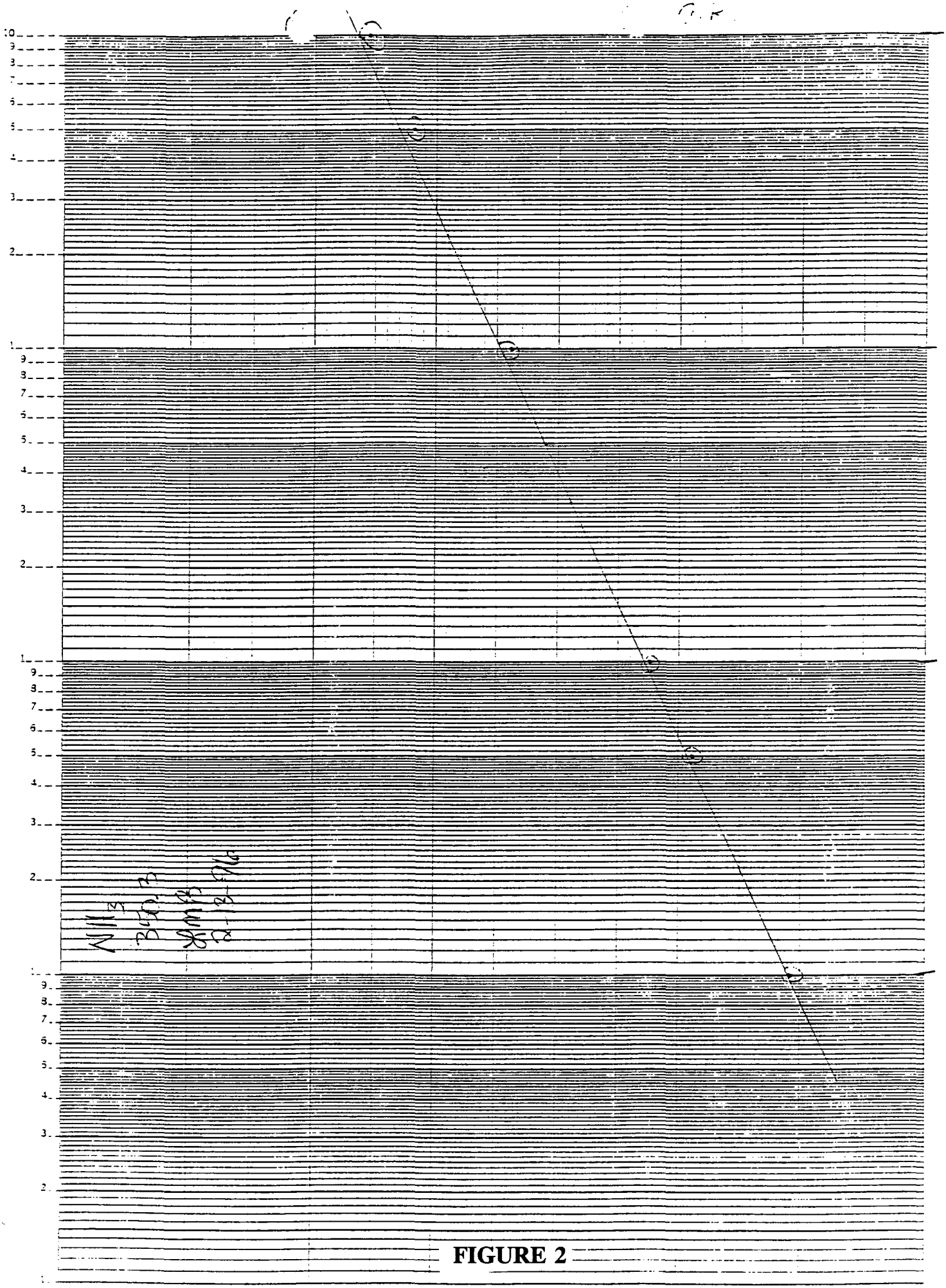

FIGURE 2